

## **REMARKS**

Claims 1-2, 4-5, and 9 stand rejected under 35 U.S.C. 102(b) as being anticipated by the publication “WebMate” (“WebMate: A Personal Agent for Browsing and Searching,” Chen et al., Proceedings of the 2<sup>nd</sup> International Conference on Autonomous Agents, May 1998, NY, USA, ACM Press, pages 132-139). Applicant traverses the rejection because the cited reference fails to disclose a document search unit which forms a keyword from the file contents transmitted from a search condition designating unit and searches similar documents from a database provided on a search side or server side where a document search unit is located.

In the Office Action (Paper No. 8), the Examiner asserts that the document search unit which forms a keyword from the file contents transmitted from the search condition designating unit of the present invention is disclosed on page 134, Col. 2, second paragraph of WebMate. More specifically, the Examiner asserts that WebMate constructs a query based on a current profile which is formed by using keywords that come from a plurality of domains, including the Webpage visited by users when the users designate them. Applicant traverses this statement of the Examiner.

WebMate forms a user’s favorite personal newspaper or the like based on the user’s profile by collecting articles or Web pages from the Internet. In the present invention, a file (e.g., a patent) forms a search request. If the file were a patent, the request would seek

other documents having keywords from the patent. This file is pre-existing, and not to be updated.

In WebMate, access to the user's Web pages is monitored and, for example, if a user adds an index such as "I like it", the index is used as a positive document and the index's HTML document is obtained. A title and a headline of a document are analyzed syntactically and nouns are extracted as keywords. The contents of the user's profile are updated by a TF-IDF method with respect to each noun, thereby automatically updating. The profile is expressed by the vector TF-IDF based on a word frequency TF according to the document analysis and a reciprocal IDF of a document frequency DF. A procedure for the updating or learning process to refine the profile is as follows: First, WebMate analyzes a morpheme with respect to the obtained HTML document, extracts nouns, and then extracts the nouns of the title and the headline. Next, the extracted nouns are used as target words and the word frequency TF and the reciprocal IDF of the document frequency are obtained to thereby form the TF-IDF vector  $V_i$ . Then, when the number of vectors of each domain and a profile is less than a specified number  $N$ , a new vector  $V_i$  is inserted into the vectors  $V = [V_1, V_2, \dots, V_n]$  of the profile.

If the number of vectors exceeds a specified number  $N$ , similarities between the new vector  $V_i$  and the present vectors  $V$  are calculated. The contents are next updated by coupling a new vector with a vector of a highest similarity. Weights of the new vectors  $V_k$  are sorted in descending order and the upper  $M$  elements are held.

A TF-IDF vector  $V_x$  is formed by obtaining the HTML documents of, for example, newspaper articles from Web pages of target domains, and similarities between the vector  $V_x$  and the vectors  $V$  of the profile are calculated. If the similarities are equal to or larger than a threshold value, then the newspaper articles are collected as a user's favorite documents and a personal newspaper is formed. However, WebMate fails to disclose or suggest a document search unit which forms a keyword from the file contents transmitted from a search condition designating unit, and searches similar documents from a database provided on a search side (*i.e.*, server side where a document search unit is located).

In contrast, the present invention provides a document search unit that forms a keyword from file contents transmitted from a search condition designating unit and searches similar documents from a database provided on a search side. In order to achieve such a result, the present invention has an interface that facilitates communication between the document search unit and the search condition designating unit. Applicant believes that WebMate does not disclose or suggest an interface capable of receiving the search request from the user like the present invention. That is, even if a document file is designated as a search condition and sent to WebMate, such a file would probably be ignored because WebMate does not have a proper interface for communication between a document search unit and a search condition designating unit.

If WebMate receives a document file designated as a search condition from a user, which corresponds to the situation where the index "I like it" has been inserted into the

Web pages viewed by a user, then WebMate recognizes the document file as a user's favorite Web page. WebMate then executes an updating or learning process for refining a profile, which is a collection of designated files. Consequently, even if a user waits for an extended time period, the user will not receive a search result of the user's desired database from WebMate.

WebMate does not search a database on the search side where a document search unit is located, but rather searches the Internet based on an updated profile that is continuously changing with time. That is, WebMate does not perform a search by using one single designated document each time WebMate searches, as in the present invention, but instead searches using an entire history of stored documents in the profile.

Alternatively, even if a document file designated as a search condition by a user is not transmitted, WebMate periodically and automatically searches for a user's favorite Web pages based on the user's profile and notifies the user of the search result. For example, WebMate can collect a user's favorite articles based on the user's profile from the Web pages of a newspaper company overnight, and then send the articles as a personal newspaper to the user the following morning. Clearly, this process of searching is contradictory to the present invention in so far as WebMate does not have a document search unit that forms a keyword from the file contents transmitted from a search designating unit because a search is not conditioned on transmitted file contents. That is, WebMate provides periodic and automatic searches, even if there is no transmission of file contents. Moreover, WebMate

does not have a database on the search side, but instead searches Web pages on the Internet. For these reasons, withdrawal of the §102 rejection of claim 1 is respectfully requested.

With respect to claim 2, Applicant traverses the rejection because WebMate fails to disclose a search condition designating unit that transmits a head file portion of the designated filed contents.

The Examiner cites WebMate page 137, lines 13-18 as the search condition designating unit. More specifically, WebMate discloses:

In WebMate agent, the *context* of the search keywords in the “relevant” web pages is used to refine the search because we think that if a user tells the system some page relevant to his sear[c]h, the context of the search keywords is more informative than the content of the page.

When the relevant pages of WebMate are specified, the system searches for a plurality of keywords used in a query expression and their context. Subsequently, with respect to each keyword, the system extracts five forward words and five backward words relative to the keyword and repeats an extracting process until all keywords in a query expression are processed. The processing routine continues until processes in which word groups are collected and stop words are deleted is completed, and a frequency of word appearance is calculated. Next, a few head words of high frequencies of appearance are used

as the latest expanded search keywords. However, in the portion of the document cited by the Examiner, WebMate fails to disclose or suggest that a search condition designating unit transmit a head file portion of the designated file contents. Rather, a method of refining the search keywords with respect to all of the contents of the relevant pages as targets is disclosed by WebMate. For this reason, claim 2 is believed allowable over the cited reference and withdrawal of the §102 rejection is respectfully requested.

With respect to claim 4, Applicant traverses the §102 rejection because WebMate fails to disclose, among other things, a search executing unit which searches similar documents by searching a search database using a keyword and notifying the search requesting source of a search result. WebMate does not disclose searching a database. The search target of WebMate is not the documents in the database, but rather the Web pages existing on the Internet, unlike the present invention which searches a database. For this reason, withdrawal of the §102 rejection is respectfully requested.

With respect to claim 5, Applicant traverses the rejection because WebMate fails to disclose a keyword forming unit that counts the number of times of appearance, showing in which documents in the index of each of the search documents stored in said document database, that each of said nouns appeared, and selects a predetermined number of upper words each having the number of times of appearance in a predetermined range, and forms a keyword.

WebMate has nouns that are extracted from HTML documents as targets, and thereafter, with respect to each noun, the word frequency TF of the document, which is being processed, and the reciprocal IDF of the document frequency for the vector corresponding to the HTML document of the profile are obtained. Then, the IT-IDF vectors are formed, and the user's favorite document based on the user's profile searches on the basis of the similarity between the IT-IDF vector and the vector of the arbitrary HTML document. WebMate does not disclose or suggest selecting words on the basis of document frequency DF, which are used as keywords.

In contrast, the present invention has nouns extracted from a document by an analysis of morphemes. The document frequencies DF indicating the indices of the search documents in the database as targets are obtained, and a predetermined number of words of high frequencies are used as keywords. The keywords are then used to search for similar documents at a search side database. Since WebMate fails to disclose or suggest selecting words on the basis of the document frequencies DF, which are used as keywords to search in a search side database, withdrawal of the §102 rejection of claim 5 is respectfully requested.

With respect to claim 9, Applicant traverses the rejection because the cited reference fails to disclose or suggest a search condition designating unit that transmits contents of a file designated by a search request picture plane of a WWW browser to a search machine of a WWW server through a network, and sends the file contents to a document search unit.

The present invention has a search requesting source as a WWW browser and a search destination is a WWW server. While WebMate uses a search engine, WebMate fails to send file documents to a document search unit. For this reason, withdrawal of the §102 rejection of claim 9 is respectfully requested.

Claim 6 stands rejected under 35 U.S.C. 103(a) as being obvious over WebMate. The arguments asserted above traversing the §102 rejection are reasserted here. Applicant also traverses the rejection because WebMate fails to disclose or suggest a range of document frequencies wherein when the number of documents in the indexes is assumed to be (N), a keyword forming unit selects the upper ten words each having a number (H) of times of appearance in a range where  $2N/3 \geq H \geq 1$ , and forms a keyword.

The Examiner asserts that the number of vectors which are inserted into the profile of WebMate is limited to N. However, the maximum number of vectors is limited in WebMate, unlike the present invention which has a range of the document frequencies DF. Accordingly, withdrawal of the §103 rejection is respectfully requested.

Claims 7-8 stand rejected under 35 U.S.C. 103(a) as being obvious over WebMate in view of the publication, "CiteSeer: An Autonomous Web Agent for Automatic Retrieval and Identification of Interesting Publications," by Bollacker et al., proceedings of the International Conference on Autonomous Agents, May 1998, ACM Press, pages 116-123 (hereinafter "CiteSeer"). The arguments asserted above traversing the §102 rejection are reasserted here. Applicant also traverses the rejection because the cited references fail to



disclose or suggest a keyword forming unit that allows property information extracted from a file received in response to a search request to be included in a keyword, which allows similar documents to be searched from a database on a search side or server side where a document unit is located.

As discussed with respect to the rejection of independent claim 1, WebMate fails to disclose or suggest searching for similar documents from a database that is located on a search side. Moreover, Applicant believes that property information which is formed into keywords that are used in a search on a search side is not taught or suggested by the cited references. For these reasons, withdrawal of the §103 rejection of claims 7-8 is respectfully requested.

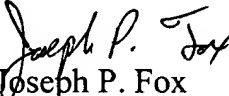
Claim 3 stands rejected under 35 U.S.C. 103(a) as being obvious over WebMate, and further in view of Eichstaedt (U.S. Patent No. 6,182,085). Since claim 3 depends upon claim 1, it necessarily includes all of the features of its associated independent claim plus other additional features. Thus, Applicant submits that the §103 rejection of claim 3 has also been overcome for the same reasons mentioned above to overcome the rejection of independent claim 1, and because Eichstaedt fails to overcome the deficiencies of the WebMate reference. Applicant respectfully requests that the §103 rejection of claim 3 also be withdrawn.

New claim 19 is claims 1, 4, and 7 combined. Applicant believes new claim 19 is allowable for the reasons recited above.

For all of the foregoing reasons, Applicant submits that this Application is in condition for allowance, which is respectfully requested. The Examiner is invited to contact the undersigned attorney if an interview would expedite prosecution.

Respectfully submitted,

GREER, BURNS & CRAIN, LTD.

By   
Joseph P. Fox  
Registration No. 41,760

July 13, 2004

300 South Wacker Drive  
Suite 2500  
Chicago, Illinois 60606  
Telephone: (312) 360-0080  
Facsimile: (312) 360-9315  
Customer No. 24978